



SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Björck, Lars  
Sjöbring, Ulf

(ii) TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF

(iii) NUMBER OF SEQUENCES: 14

(iv) CORRESPONDENCE ADDRESS:

- (A) ADDRESSEE: SEED and BERRY LLP  
(B) STREET: 6300 Columbia Center, 701 Fifth Avenue  
(C) CITY: Seattle  
(D) STATE: Washington  
(E) COUNTRY: USA  
(F) ZIP: 98104-7092

(v) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER: US 08/325,278  
(B) FILING DATE: 26-OCT-1996  
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: McMasters, David D.  
(B) REGISTRATION NUMBER: 33,963  
(C) REFERENCE/DOCKET NUMBER: 450023.401

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: (206) 622-4900  
(B) TELEFAX: (206) 682-6031

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 305 amino acids

(B) TYPE: amino acid  
(C) STRANDEDNESS: unknown  
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Escherichia coli LE392/pHDL, DSM 7054

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Ala	Val	Glu	Asn	Lys	Glu	Glu	Thr	Pro	Glu	Thr	Pro	Glu	Thr	Asp	Ser
1				5				10						15	
Glu	Glu	Glu	Val	Thr	Ile	Lys	Ala	Asn	Leu	Ile	Phe	Ala	Asn	Gly	Ser
			20					25					30		
Thr	Gln	Thr	Ala	Glu	Phe	Lys	Gly	Thr	Phe	Glu	Lys	Ala	Thr	Ser	Glu
			35				40					45			
Ala	Tyr	Ala	Tyr	Ala	Asp	Thr	Leu	Lys	Lys	Asp	Asn	Gly	Glu	Tyr	Thr
	50					55					60				
Val	Asp	Val	Ala	Asp	Lys	Gly	Tyr	Thr	Leu	Asn	Ile	Lys	Phe	Ala	Gly
65					70				75						80
Lys	Glu	Lys	Thr	Pro	Glu	Glu	Pro	Lys	Glu	Glu	Val	Thr	Ile	Lys	Ala
				85					90					95	
Asn	Leu	Ile	Tyr	Ala	Asp	Gly	Lys	Thr	Gln	Thr	Ala	Glu	Phe	Lys	Gly
			100					105					110		
Thr	Phe	Glu	Glu	Ala	Thr	Ala	Glu	Ala	Tyr	Arg	Tyr	Ala	Asp	Ala	Leu
			115				120					125			
Lys	Lys	Asp	Asn	Gly	Glu	Tyr	Thr	Val	Asp	Val	Ala	Asp	Lys	Gly	Tyr
	130					135					140				
Thr	Leu	Asn	Ile	Lys	Phe	Ala	Gly	Lys	Glu	Lys	Thr	Pro	Glu	Glu	Pro
145					150				155						160

Lys Glu Glu Val Thr Ile Lys Ala Asn Leu Ile Tyr Ala Asp Gly Lys  
 165 170 175  
 Thr Gln Thr Ala Glu Phe Lys Gly Thr Phe Glu Glu Ala Thr Ala Glu  
 180 185 190  
 Ala Tyr Arg Tyr Ala Asp Leu Leu Ala Lys Glu Asn Gly Lys Tyr Thr  
 195 200 205  
 Val Asp Val Ala Asp Lys Gly Tyr Thr Leu Asn Ile Lys Phe Ala Gly  
 210 215 220  
 Lys Glu Lys Thr Pro Glu Glu Pro Lys Glu Glu Val Thr Ile Lys Ala  
 225 230 235 240  
 Asn Leu Ile Tyr Ala Asp Gly Lys Thr Gln Thr Ala Glu Phe Lys Gly  
 245 250 255  
 Thr Phe Ala Glu Ala Thr Ala Glu Ala Tyr Arg Tyr Ala Asp Leu Leu  
 260 265 270  
 Ala Lys Glu Asn Gly Lys Tyr Thr Ala Asp Leu Glu Asp Gly Gly Tyr  
 275 280 285  
 Thr Ile Asn Ile Arg Phe Ala Gly Lys Lys Val Asp Glu Lys Pro Glu  
 290 295 300  
 Glu  
 305

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 921 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Escherichia coli LE392/pHDL, DSM 7054

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GCGGTAGAAA ATAAAGAAGA AACACCAGAA ACACCAGAAA CTGATTCAGA AGAAGAAGTA	60
ACAATCAAAG CTAACCTAAT CTTTGCAAAT GGAAGCACAC AACTGCAGA ATTCAAAGGA	120
ACATTTGAAA AAGCAACATC AGAAGCTTAT GCGTATGCAG ATACTTTGAA GAAAGACAAT	180
GGAGAATATA CTGTAGATGT TGCAGATAAA GGTTATACTT TAAATATTAA ATTTGCTGGA	240
AAAGAAAAAA CACCAGAAGA ACCAAAAGAA GAAGTTACTA TTAAAGCAAA CTTAATCTAT	300
GCAGATGGAA AACACAAAC AGCAGAATTC AAAGGAACAT TTGAAGAAGC AACAGCAGAA	360
GCATACAGAT ATGCAGATGC ATTAAAGAAG GACAATGGAG AATATACAGT AGACGTTGCA	420
GATAAAGGTT ATACTTTAAA TATTAAATTT GCTGGAAAAG AAAAAACACC AGAAGAACCA	480
AAAGAAGAAG TTACTATTAA AGCAAACTTA ATCTATGCAG ATGGAAAAAC ACAAACAGCA	540
GAATTCAAAG GAACATTTGA AGAAGCAACA GCAGAAGCAT ACAGATATGC TGACTTATTA	600
GCAAAAGAAA ATGGTAAATA TACAGTAGAC GTTGCAGATA AAGGTTATAC TTAAATATT	660
AAATTTGCTG GAAAAGAAAA AACACCAGAA GAACCAAAAG AAGAAGTTAC TATTAAAGCA	720
AACTTAATCT ATGCAGATGG AAAAActCAA ACAGCAGAGT TCAAAGGAAC ATTTGCAGAA	780
GCAACAGCAG AAGCATACAG ATACGCTGAC TTATTAGCAA AAGAAAATGG TAAATATACA	840
GCAGACTTAG AAGATGGTGG ATACACTATT AATATTAGAT TTGCAGGTAA GAAAGTTGAC	900
GAAAAACCAG AAGAATAATA A	921

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 434 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Escherichia coli LE392/pHDLG, DSM 7055

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Ala	Val	Glu	Asn	Lys	Glu	Glu	Thr	Pro	Glu	Thr	Pro	Glu	Thr	Asp	Ser
1				5				10						15	
Glu	Glu	Glu	Val	Thr	Ile	Lys	Ala	Asn	Leu	Ile	Phe	Ala	Asn	Gly	Ser
			20					25					30		
Thr	Gln	Thr	Ala	Glu	Phe	Lys	Gly	Thr	Phe	Glu	Lys	Ala	Thr	Ser	Glu
		35					40					45			
Ala	Tyr	Ala	Tyr	Ala	Asp	Thr	Leu	Lys	Lys	Asp	Asn	Gly	Glu	Tyr	Thr
	50					55					60				
Val	Asp	Val	Ala	Asp	Lys	Gly	Tyr	Thr	Leu	Asn	Ile	Lys	Phe	Ala	Gly
65					70					75					80
Lys	Glu	Lys	Thr	Pro	Glu	Glu	Pro	Lys	Glu	Glu	Val	Thr	Ile	Lys	Ala
				85					90					95	
Asn	Leu	Ile	Tyr	Ala	Asp	Gly	Lys	Thr	Gln	Thr	Ala	Glu	Phe	Lys	Gly
			100					105					110		
Thr	Phe	Glu	Glu	Ala	Thr	Ala	Glu	Ala	Tyr	Arg	Tyr	Ala	Asp	Ala	Leu
		115					120					125			
Lys	Lys	Asp	Asn	Gly	Glu	Tyr	Thr	Val	Asp	Val	Ala	Asp	Lys	Gly	Tyr
	130					135					140				
Thr	Leu	Asn	Ile	Lys	Phe	Ala	Gly	Lys	Glu	Lys	Thr	Pro	Glu	Glu	Pro
145					150				155						160
Lys	Glu	Glu	Val	Thr	Ile	Lys	Ala	Asn	Leu	Ile	Tyr	Ala	Asp	Gly	Lys
				165					170					175	
Thr	Gln	Thr	Ala	Glu	Phe	Lys	Gly	Thr	Phe	Glu	Glu	Ala	Thr	Ala	Glu
			180					185					190		

Ala	Tyr	Arg	Tyr	Ala	Asp	Leu	Leu	Ala	Lys	Glu	Asn	Gly	Lys	Tyr	Thr	195	200	205
Val	Asp	Val	Ala	Asp	Lys	Gly	Tyr	Thr	Leu	Asn	Ile	Lys	Phe	Ala	Gly	210	215	220
Lys	Glu	Lys	Thr	Pro	Glu	Glu	Pro	Lys	Glu	Glu	Val	Thr	Ile	Lys	Ala	225	230	235
Asn	Leu	Ile	Tyr	Ala	Asp	Gly	Lys	Thr	Gln	Thr	Ala	Glu	Phe	Lys	Gly	245	250	255
Thr	Phe	Ala	Glu	Ala	Thr	Ala	Glu	Ala	Tyr	Arg	Tyr	Ala	Asp	Leu	Leu	260	265	270
Ala	Lys	Glu	Asn	Gly	Lys	Tyr	Thr	Ala	Asp	Leu	Glu	Asp	Gly	Gly	Tyr	275	280	285
Thr	Ile	Asn	Ile	Arg	Phe	Ala	Gly	Lys	Lys	Val	Asp	Glu	Lys	Pro	Glu	290	295	300
Glu	Pro	Met	Asp	Thr	Tyr	Lys	Leu	Ile	Leu	Asn	Gly	Lys	Thr	Leu	Lys	305	310	315
Gly	Glu	Thr	Thr	Thr	Glu	Ala	Val	Asp	Ala	Ala	Thr	Ala	Glu	Lys	Val	325	330	335
Phe	Lys	Gln	Tyr	Ala	Asn	Asp	Asn	Gly	Val	Asp	Gly	Glu	Trp	Thr	Tyr	340	345	350
Asp	Asp	Ala	Thr	Lys	Thr	Phe	Thr	Val	Thr	Glu	Lys	Pro	Glu	Val	Ile	355	360	365
Asp	Ala	Ser	Glu	Leu	Thr	Pro	Ala	Val	Thr	Thr	Tyr	Lys	Leu	Val	Ile	370	375	380
Asn	Gly	Lys	Thr	Leu	Lys	Gly	Glu	Thr	Thr	Thr	Lys	Ala	Val	Asp	Ala	385	390	395
Glu	Thr	Ala	Glu	Lys	Ala	Phe	Lys	Gln	Tyr	Ala	Asn	Asp	Asn	Gly	Val	405	410	415
Asp	Gly	Val	Trp	Thr	Tyr	Asp	Asp	Ala	Thr	Lys	Thr	Phe	Thr	Val	Thr	420	425	430

Glu Met

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1308 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Escherichia coli L392/pHDLG, DSM 7055

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GCGGTAGAAA ATAAAGAAGA AACACCAGAA ACACCAGAAA CTGATTCAGA AGAAGAAGTA	60
ACAATCAAAG CTAACCTAAT CTTTGCAAAT GGAAGCACAC AAAGTGCAGA ATTCAAAGGA	120
ACATTTGAAA AAGCAACATC AGAAGCTTAT GCGTATGCAG ATACTTTGAA GAAAGACAAT	180
GGAGAATATA CTGTAGATGT TGCAGATAAA GGTTATACTT TAAATATTAA ATTTGCTGGA	240
AAAGAAAAAA CACCAGAAGA ACCAAAAGAA GAAGTTACTA TTAAAGCAAA CTTAATCTAT	300
GCAGATGGAA AAACACAAAC AGCAGAATTC AAAGGAACAT TTGAAGAAGC AACAGCAGAA	360
GCATACAGAT ATGCAGATGC ATTAAGAAG GACAATGGAG AATATACAGT AGACGTTGCA	420
GATAAAGGTT ATACTTTAAA TATTAAATTT GCTGGAAAAG AAAAAACACC AGAAGAACCA	480
AAAGAAGAAG TTACTATTAA AGCAAACTTA ATCTATGCAG ATGGAAAAAC ACAAACAGCA	540
GAATTCAAAG GAACATTTGA AGAAGCAACA GCAGAAGCAT ACAGATATGC TGAATTATTA	600
GCAAAAGAAA ATGGTAAATA TACAGTAGAC GTTGCAGATA AAGGTTATAC TTAAATATT	660

AAATTTGCTG	GAAAAGAAAA	AACACCAGAA	GAACCAAAAG	AAGAAGTTAC	TATTAAAGCA	720
AACTTAATCT	ATGCAGATGG	AAAAACTCAA	ACAGCAGAGT	TCAAAGGAAC	ATTTGCAGAA	780
GCAACAGCAG	AAGCATACAG	ATACGCTGAC	TTATTAGCAA	AAGAAAATGG	TAAATATACA	840
GCAGACTTAG	AAGATGGTGG	ATACACTATT	AATATTAGAT	TTGCAGGTAA	GAAAGTTGAC	900
GAAAAACCAG	AAGAACCCAT	GGACACTTAC	AAATTAATCC	TTAATGGTAA	AACATTGAAA	960
GGCGAAACAA	CTACTGAAGC	TGTTGATGCT	GCTACTGCAG	AAAAAGTCTT	CAAACAATAC	1020
GCTAACGACA	ACGGTGTTGA	CGGTGAATGG	ACTTACGACG	ATGCGACTAA	GACCTTTACA	1080
GTTACTGAAA	AACCAGAAAGT	GATCGATGCG	TCTGAATTAA	CACCAGCCGT	GACAACTTAC	1140
AACTTGTTA	TTAATGGTAA	AACATTGAAA	GGCGAAACAA	CTACTAAAGC	AGTAGACGCA	1200
GAAACTGCAG	AAAAAGCCTT	CAAACAATAC	GCTAACGACA	ACGGTGTTGA	TGGTGTTTGG	1260
ACTTATGATG	ATGCGACTAA	GACCTTTACG	GTAAGTAAAA	TGTAATAA		1308

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1332 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: unknown

## (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(ix) FEATURE:

- (A) NAME/KEY: CDS  
(B) LOCATION: 1..1329

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AAC GGT GAT GGT AAT CCT AGG GAA GTT ATA GAA GAT CTT GCA GCA AAC 48  
Asn Gly Asp Gly Asn Pro Arg Glu Val Ile Glu Asp Leu Ala Ala Asn  
1 5 10 15



AAT CCC GCA ATA CAA AAT ATA CGT TTA CGT CAC GAA AAC AAG GAC TTA Asn Pro Ala Ile Gln Asn Ile Arg Leu Arg His Glu Asn Lys Asp Leu	96
20 25 30	
AAA GCG AGA TTA GAG AAT GCA ATG GAA GTT GCA GGA AGA GAT TTT AAG Lys Ala Arg Leu Glu Asn Ala Met Glu Val Ala Gly Arg Asp Phe Lys	144
35 40 45	
AGA GCT GAA GAA CTT GAA AAA GCA AAA CAA GCC TTA GAA GAC CAG CGT Arg Ala Glu Glu Leu Glu Lys Ala Lys Gln Ala Leu Glu Asp Gln Arg	192
50 55 60	
AAA GAT TTA GAA ACT AAA TTA AAA GAA CTA CAA CAA GAC TAT GAC TTA Lys Asp Leu Glu Thr Lys Leu Lys Glu Leu Gln Gln Asp Tyr Asp Leu	240
65 70 75 80	
GCA AAG GAA TCA ACA AGT TGG GAT AGA CAA AGA CTT GAA AAA GAG TTA Ala Lys Glu Ser Thr Ser Trp Asp Arg Gln Arg Leu Glu Lys Glu Leu	288
85 90 95	
GAA GAG AAA AAG GAA GCT CTT GAA TTA GCG ATA GAC CAG GCA AGT CGG Glu Glu Lys Lys Glu Ala Leu Glu Leu Ala Ile Asp Gln Ala Ser Arg	336
100 105 110	
GAC TAC CAT AGA GCT ACC GCT TTA GAA AAA GAG TTA GAA GAG AAA AAG Asp Tyr His Arg Ala Thr Ala Leu Glu Lys Glu Leu Glu Glu Lys Lys	384
115 120 125	
AAA GCT CTT GAA TTA GCG ATA GAC CAA GCG AGT CAG GAC TAT AAT AGA Lys Ala Leu Glu Leu Ala Ile Asp Gln Ala Ser Gln Asp Tyr Asn Arg	432
130 135 140	
GCT AAC GTC TTA GAA AAA GAG TTA GAA ACG ATT ACT AGA GAA CAA GAG Ala Asn Val Leu Glu Lys Glu Leu Glu Thr Ile Thr Arg Glu Gln Glu	480
145 150 155 160	
ATT AAT CGT AAT CTT TTA GGC AAT GCA AAA CTT GAA CTT GAT CAA CTT Ile Asn Arg Asn Leu Leu Gly Asn Ala Lys Leu Glu Leu Asp Gln Leu	528
165 170 175	
TCA TCT GAA AAA GAG CAG CTA ACG ATC GAA AAA GCA AAA CTT GAG GAA Ser Ser Glu Lys Glu Gln Leu Thr Ile Glu Lys Ala Lys Leu Glu Glu	576
180 185 190	

GAA AAA CAA ATC TCA GAC GCA AGT CGT CAA AGC CTT CGT CGT GAC TTG Glu Lys Gln Ile Ser Asp Ala Ser Arg Gln Ser Leu Arg Arg Asp Leu 195 200 205	624
GAC GCA TCA CGT GAA GCT AAG AAA CAG GTT GAA AAA GAT TTA GCA AAC Asp Ala Ser Arg Glu Ala Lys Lys Gln Val Glu Lys Asp Leu Ala Asn 210 215 220	672
TTG ACT GCT GAA CTT GAT AAG GTT AAA GAA GAC AAA CAA ATC TCA GAC Leu Thr Ala Glu Leu Asp Lys Val Lys Glu Asp Lys Gln Ile Ser Asp 225 230 235 240	720
GCA AGC CGT CAA CGG CTT CGC CGT GAC TTG GAC GCA TCA CGT GAA GCT Ala Ser Arg Gln Arg Leu Arg Arg Asp Leu Asp Ala Ser Arg Glu Ala 245 250 255	768
AAG AAA CAG GTT GAA AAA GAT TTA GCA AAC TTG ACT GCT GAA CTT GAT Lys Lys Gln Val Glu Lys Asp Leu Ala Asn Leu Thr Ala Glu Leu Asp 260 265 270	816
AAG GTT AAA GAA GAA AAA CAA ATC TCA GAC GCA AGC CGT CAA CGG CTT Lys Val Lys Glu Glu Lys Gln Ile Ser Asp Ala Ser Arg Gln Arg Leu 275 280 285	864
CGC CGT GAC TTG GAC GCA TCA CGT GAA GCT AAG AAA CAA GTT GAA AAA Arg Arg Asp Leu Asp Ala Ser Arg Glu Ala Lys Lys Gln Val Glu Lys 290 295 300	912
GCT TTA GAA GAA GCA AAC AGC AAA TTA GCT GCT CTT GAA AAA CTT AAC Ala Leu Glu Glu Ala Asn Ser Lys Leu Ala Ala Leu Glu Lys Leu Asn 305 310 315 320	960
AAA GAG CTT GAA GAA AGC AAG AAA TTA ACA GAA AAA GAA AAA GCT GAA Lys Glu Leu Glu Glu Ser Lys Lys Leu Thr Glu Lys Glu Lys Ala Glu 325 330 335	1008
CTA CAA GCA AAA CTT GAA GCA GAA GCA AAA GCA CTC AAA GAA CAA TTA Leu Gln Ala Lys Leu Glu Ala Glu Ala Lys Ala Leu Lys Glu Gln Leu 340 345 350	1056
GCG AAA CAA GCT GAA GAA CTC GCA AAA CTA AGA GCT GGA AAA GCA TCA Ala Lys Gln Ala Glu Glu Leu Ala Lys Leu Arg Ala Gly Lys Ala Ser 355 360 365	1104
GAC TCA CAA ACC CCT GAT ACA AAA CCA GGA AAC AAA GCT CTT CCA GGT	1152

Asp	Ser	Gln	Thr	Pro	Asp	Thr	Lys	Pro	Gly	Asn	Lys	Val	Leu	Pro	Gly	
370						375					380					
AAA	GGT	CAA	GCA	CCA	CAA	GCA	GGT	ACA	AAA	CCT	AAC	CAA	AAC	AAA	GCA	1200
Lys	Gly	Gln	Ala	Pro	Gln	Ala	Gly	Thr	Lys	Pro	Asn	Gln	Asn	Lys	Ala	
385					390					395					400	
CCA	ATG	AAG	GAA	ACT	AAG	AGA	CAG	TTA	CCA	TCA	ACA	GGT	GAA	ACA	GCT	1248
Pro	Met	Lys	Glu	Thr	Lys	Arg	Gln	Leu	Pro	Ser	Thr	Gly	Glu	Thr	Ala	
				405					410					415		
AAC	CCA	TTC	TTC	ACA	GCG	GCA	CGC	GTT	ACT	GTT	ATG	GCA	ACA	GCT	GGA	1296
Asn	Pro	Phe	Phe	Thr	Ala	Ala	Arg	Val	Thr	Val	Met	Ala	Thr	Ala	Gly	
			420					425					430			
GTA	GCA	GCA	GTT	GTA	AAA	CGC	AAA	GAA	GAA	AAC	TAA					1332
Val	Ala	Ala	Val	Val	Lys	Arg	Lys	Glu	Glu	Asn						
			435				440									

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 443 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Asn	Gly	Asp	Gly	Asn	Pro	Arg	Glu	Val	Ile	Glu	Asp	Leu	Ala	Ala	Asn	
1				5					10					15		
Asn	Pro	Ala	Ile	Gln	Asn	Ile	Arg	Leu	Arg	His	Glu	Asn	Lys	Asp	Leu	
			20					25					30			
Lys	Ala	Arg	Leu	Glu	Asn	Ala	Met	Glu	Val	Ala	Gly	Arg	Asp	Phe	Lys	
		35					40					45				
Arg	Ala	Glu	Glu	Leu	Glu	Lys	Ala	Lys	Gln	Ala	Leu	Glu	Asp	Gln	Arg	
	50					55					60					
Lys	Asp	Leu	Glu	Thr	Lys	Leu	Lys	Glu	Leu	Gln	Gln	Asp	Tyr	Asp	Leu	
65					70					75					80	

Ala Lys Glu Ser Thr Ser Trp Asp Arg Gln Arg Leu Glu Lys Glu Leu  
85 90 95

Glu Glu Lys Lys Glu Ala Leu Glu Leu Ala Ile Asp Gln Ala Ser Arg  
100 105 110

Asp Tyr His Arg Ala Thr Ala Leu Glu Lys Glu Leu Glu Glu Lys Lys  
115 120 125

Lys Ala Leu Glu Leu Ala Ile Asp Gln Ala Ser Gln Asp Tyr Asn Arg  
130 135 140

Ala Asn Val Leu Glu Lys Glu Leu Glu Thr Ile Thr Arg Glu Gln Glu  
145 150 155 160

Ile Asn Arg Asn Leu Leu Gly Asn Ala Lys Leu Glu Leu Asp Gln Leu  
165 170 175

Ser Ser Glu Lys Glu Gln Leu Thr Ile Glu Lys Ala Lys Leu Glu Glu  
180 185 190

Glu Lys Gln Ile Ser Asp Ala Ser Arg Gln Ser Leu Arg Arg Asp Leu  
195 200 205

Asp Ala Ser Arg Glu Ala Lys Lys Gln Val Glu Lys Asp Leu Ala Asn  
210 215 220

Leu Thr Ala Glu Leu Asp Lys Val Lys Glu Asp Lys Gln Ile Ser Asp  
225 230 235 240

Ala Ser Arg Gln Arg Leu Arg Arg Asp Leu Asp Ala Ser Arg Glu Ala  
245 250 255

Lys Lys Gln Val Glu Lys Asp Leu Ala Asn Leu Thr Ala Glu Leu Asp  
260 265 270

Lys Val Lys Glu Glu Lys Gln Ile Ser Asp Ala Ser Arg Gln Arg Leu  
275 280 285

Arg Arg Asp Leu Asp Ala Ser Arg Glu Ala Lys Lys Gln Val Glu Lys  
290 295 300

Ala Leu Glu Glu Ala Asn Ser Lys Leu Ala Ala Leu Glu Lys Leu Asn  
305 310 315 320

Lys Glu Leu Glu Glu Ser Lys Lys Leu Thr Glu Lys Glu Lys Ala Glu  
 325 330 335  
 Leu Gln Ala Lys Leu Glu Ala Glu Ala Lys Ala Leu Lys Glu Gln Leu  
 340 345 350  
 Ala Lys Gln Ala Glu Glu Leu Ala Lys Leu Arg Ala Gly Lys Ala Ser  
 355 360 365  
 Asp Ser Gln Thr Pro Asp Thr Lys Pro Gly Asn Lys Ala Val Pro Gly  
 370 375 380  
 Lys Gly Gln Ala Pro Gln Ala Gly Thr Lys Pro Asn Gln Asn Lys Ala  
 385 390 395 400  
 Pro Met Lys Glu Thr Lys Arg Gln Leu Pro Ser Thr Gly Glu Thr Ala  
 405 410 415  
 Asn Pro Phe Phe Thr Ala Ala Arg Val Thr Val Met Ala Thr Ala Gly  
 420 425 430  
 Val Ala Ala Val Val Lys Arg Lys Glu Glu Asn  
 435 440

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 44 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GCTCAGGCGG CGCCGGTAGA AAATAAAGAA GAAACACCAG AAAC

44

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 9 amino acids

- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Val Glu Asn Lys Glu Glu Thr Pro Glu  
1 5

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 47 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CAGCAGCAGG ATTCTTATTA TTCTTCTGGT TTTTCGTCAA CTTTCTT

47

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 44 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CAGCAGCAGC CATGGGTTCT TCTGGTTTTT CGTCAACTTT CTTA

44

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 34 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGCCATGGAC ACTTACAAAT TAATCCTTAA TGGT

34

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 10 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear

*Sub H1*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met Asp Thr Tyr Lys Leu Ile Leu Asn Gly  
1                      5                      10

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 42 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CAGGTCGACT TATTACATTT CAGTTACCGT AAAGGTCTTA GT

42

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 152 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

Sub  
H,

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

AAGCTTAAGG AGGTTAATCG ATGAAAAAAA CTGCTATCGC TATCGCTGTT GCTCTGGCTG 60

GTTTCGCTAC TGTTGCTCAG GCGGCGCCGA GATCTAAACA GGAATTCGAG CTCGGTACCC 120

GGGGATCCTC TAGAGCTGAC CTGCAGGCAT GC 152